

## **ICSU Panel on World Data Centers**

Bremen, Germany

February 1-2, 2005

Meeting Summary

### **Attendance**

Panel: Ferris Webster (chair), Tohru Araki, Mike Chinnery, David Clark, Alexei Gvishiani, Shengli Huang (for Sun Jiulin), Anne Linn, Gennady Sobolev, and Yuri Tyupkin.

Guests: Carthage Smith (ICSU), Michael Diepenbroek (WDC MARE), Manfred Reinke (Alfred Wegener Institute), Hannes Grobbe (Alfred Wegener Institute), Uwe Schindler (WDC MARE)

- Linn will ask for the names and email addresses of Germans who participated in the meeting.

### **ICSU Priority Area Assessment on Data and Information**

The International Council for Science (ICSU) is commissioning a number of priority area assessments (PAA) to help it develop an overall strategic vision. The PAA on data and information addresses a number of issues that are relevant to the WDCs. What are ICSU's data needs? Can existing ICSU structures respond? The report points out the need for professional data management, long-term (century) thinking, and coordination of ICSU and other programs. A long-term strategic framework for data and information will be developed by an ad hoc committee, which will organize fora every few years. Representatives from each of the major ICSU data bodies will be on the ad hoc committee. The WDC Panel should take the initiative to make changes in the WDC system, rather than waiting for direction from the ad hoc committee. ICSU cannot offer money to make these changes, but the report and ICSU endorsement should help the Panel help itself.

Gvishiani asked how ICSU saw the roles of CODATA and the WDC Panel. Smith said that there is no feeling that the two bodies, which do different things, should merge, although they should work together on issues of common interest, such as access to data. Both the WDC Panel and CODATA should take a more strategic approach to their varied activities. The WDC system needs better coordination, greater use of standards, and some WDCs need to evolve to better meet user needs. The WDC modernization report commissioned by the Panel acknowledges these issues.

Webster said that the WDC system does not contain all data needed to serve the needs of users. For example, the need for a hydrology archive was identified at a WDC-IGBP workshop, but attempts to establish a WDC for hydrology have failed. The WDCs need to maintain a dialog with scientific data programs and observing systems. Many of these expect that the data will end up at the WDCs, even though that role is not formally recognized in their strategic plans.

Diepenbroek noted that his institute has had considerable success getting money from national and European science programs for data management. But the reputation of the WDCs is important in this context. If users can't get data from the center, the programs won't ask them to

do their data management. Linn noted that the Panel always encourages this approach, but in practice, the WDCs are rarely involved in the planning stages and they are often expected to take the data without any funding.

Webster noted that an immediate step that can be taken to improve the WDC system is to develop a directory of holdings. The Panel has been trying to get the WDCs to do this for several years, but these efforts have failed thus far. Less than half of the WDCs have provided a simple html list of holdings and some think it's not necessary to create a catalog at all. There are 53 data centers, but are they all functioning? Can users find data from all of them or even contact them?

The Panel decided that its top priority for responding to the PAA on data and information is to strengthen the WDC system through a certification process.

### **Certification of WDCs**

The issue of certifying the quality of WDCs has been a topic of discussion for several years. The Panel decided that levels of certification should be based on the significance of the holdings, technology (i.e., have a directory), adherence to the WDC full and open data policy, and user needs (if users feel they can't get useful data from the center, they won't bother with the WDC system). The three levels of certification are:

#### World Data Center

1. Has data holdings that are significant.
2. Provides data on a nondiscriminatory basis, without restriction, for no more than the cost of reproduction.
3. Maintains a web-accessible directory of holdings that adhere to national or international standards.
4. Provides a means for users to find and download data online.
5. Reports to the WDC Panel as requested.

#### Associate Center

1. Provides data on a nondiscriminatory basis, without restriction, for no more than the cost of reproduction.
2. Maintains a directory of holdings.
3. Provides a means for users to obtain data.
4. Reports to the WDC Panel as requested.

#### Centers Under Development

Centers working to achieve associate or world data center status (e.g., developing countries). There are no rules thus far for these centers, which tend to be in developing countries.

The certification process may leave only those centers that still want to be WDCs. Proposed new centers would have to apply for full WDC status. Associate status would be available only for existing data centers that need encouragement to improve. Associates will strive to become full WDCs because (1) data providers may not want to send their data to an associate center, (2) it

may be easier to obtain funding with the WDC title, and (3) peer pressure. The goal is to have a WDC system composed of full WDC members.

Some of the China WDCs might become associates because they have to pay for users to download data and consequently don't offer data via the Internet. However, the centers could pass that charge to users and/or they could accept the offer of sister centers in other countries to mirror their holdings and thus retain WDC status.

There was some discussion about having time limits for associate centers. For example, should associates who don't advance to WDC status within 3 years take a step back? The Panel will consider this issue after the certification process concludes.

Unlike the other specific criteria, assessing the significance of the holdings (full WDC criteria) is subjective. How can the Panel make the determination without a costly review? The number of requests and users can be used as a proxy of data importance, but these numbers can be manipulated easily. Smith suggested asking relevant ICSU bodies whether the holdings in question make a contribution. Linn noted that you can't evaluate the significance of the holdings without knowing what they are. She proposed certifying the centers in two stages: (1) obtain a catalog of holdings and (2) peer review of the significance of the holdings.

The Panel decided to begin the WDC certification process using the levels and criteria defined above. Recertification should be done on a staggered basis every 10 years. The China centers will be certified as part of the China WDC review in 2005.

- Webster will draft a letter to the WDCs explaining the certification goals and process and seeking their opinion on their level of certification. ICSU and the relevant ICSU national member will be copied on the letter. The centers will be given up to two reminder letters, then nonresponding centers will be closed using procedures established by the WDC Panel.

### **International Polar Year (IPY)**

Smith suggested that the IPY offered an opportunity to address some of the important challenges facing the WDC system. IPY would lever new funding and offer an opportunity for the data centers to provide leadership. The IPY data plan is currently very preliminary—what is the WDC role? Linn asked if ICSU would entertain a proposal for rescuing polar data sets to help interpret the IPY data sets. Smith was noncommittal about ICSU funding, but suggested talking with NSF and the IPY project office about submitting an IPY project proposal. ICSU itself doesn't have any project funding for IPY.

Clark noted that the WDC for Glaciology, Boulder, has submitted a letter of intent for IPY activities. Smith said that a joint proposal from several WDCs should also be considered.

- Clark will discuss with Florence Fetterer the potential for involving multiple WDCs in specific IPY proposals.

## **Global Earth Observing System of Systems (GEOSS)**

More than 30 countries have agreed to develop a comprehensive, coordinated, and sustained Earth observing system to collect and disseminate data, information, and models to stakeholders and decision makers. The system is in the planning stages and data centers are not mentioned in the planning documents. Smith said the WDC system should develop a plan for getting involved.

- Webster and Linn will talk with Greg Withee (U.S. Group on Earth Observations [GEO] chair) about WDC involvement in GEOSS.
- Clark will approach Asbindu Singh (international GEO data panel) about including the WDCs in the GEOSS 10-year plan.

## **Electronic Geophysical Year (eGY)**

The eGY initiative focuses on geophysical data and e-science as part of the 50th anniversary of the International Geophysical Year (IGY). It is in the planning stages and has yet to establish clear goals or scientific objectives or to reach out to other disciplines. The main focus to date is on virtual observatories and solar-terrestrial physics, probably because the initiative came out of the International Association of Geomagnetism and Aeronomy (IAGA). Suggestions from WDC representatives that the eGY focus on IPY have not been well received. However, if the initiative goes forward, the WDCs must play a role. The Panel has participated in all eGY planning meetings held thus far.

Smith said the IPY and GEOSS should perhaps be higher priorities for the WDC system, although the Panel should continue to keep an active brief on eGY.

- The Panel decided to continue participating in eGY planning meetings and become more involved as appropriate.
- Webster will notify the eGY secretary (Bill Peterson) that Clark and Chinnery will attend the next eGY meeting, which will be held in Boulder.

## **European WDC Consortium**

Diepenbroek has been trying to create a science-driven consortium of data centers to increase the political weight of WDCs in Europe and to increase funding (consortiums fare better in proposals). However, an attempt to contact the European WDCs resulted in only three replies (the German WDCs and the WDC for Glaciology, Cambridge). Further attempts will not be made until the WDC system is streamlined through the certification process. In the meantime, WDC MARE will continue to seek involvement in international projects, bringing in other WDCs as appropriate, and will go to Brussels to try to increase funding. The coordinated German efforts might succeed better if supported by a letter from the WDC Panel.

Webster asked the staff of the WDC MARE if they would consider establishing a WDC Europe coordination office. The creation of three new centers in Germany provides is an opportunity for

a new beginning in Europe. The WDC MARE staff were too busy, but they agreed to reconsider the idea after the certification process is completed.

- The Panel will talk with WDC MARE staff about creating a WDC Europe coordination office after the certification process has been completed.

### **China WDC Review**

A few years ago the Panel decided to review all of the China WDCs because of concerns about data access. Many centers are not connected to the Internet and some may not be disseminating data. However, it has been difficult to contact the centers and the WDC China coordination office to arrange a review. Many centers and the coordination office do not respond to email, fax, or letters. A Chinese delegation visited the United States in December and the head (Guo Huandong, deputy secretary of the Chinese Academy of Sciences) invited the WDC Panel to review all of the China WDCs at the Panel's convenience.

The audience for the review should not be the data centers, but higher-level administrators. Ideally, these Chinese administrators would make the decision to reduce the number of centers, if necessary. Smith noted that the ICSU general assembly, which is being held in Suzhou in October 2005, has attracted high level ministerial interest in China. This event and the ICSU strategic plan might be used as a stimulus to help move the review along.

The China WDCs cover a wide range of disciplines. However, the Panel can afford to send only a small review team to China. (The country hosting the data centers usually pays significant expenses associated with the review.) The review team will consist of Chinese scientists not connected with the WDCs and 3-4 individuals named by the Panel. Sun Honglie has provided a list of Chinese scientists. Possible Panel members or representatives include Tohru Araki, Roger Barry (director of the WDC for Glaciology, Boulder), and Chinese scientists living in the United States or Europe. The latter would be able to help the review team communicate with the WDCs before the review and interpret Chinese presentations at the review.

In addition to the WDC China review, the Chinese have submitted proposals to create a WDC Office for Capacity Building to train developing countries in data systems, and a WDC for Ecosystems in Mountains and Plateaus.

- The Panel decided to review the China WDCs and the proposals for a new office and WDC simultaneously. The criteria for evaluating new centers will be the standard ones established by the Panel. The criteria for evaluating existing WDCs will be the minimum requirements for WDCs laid out in the certification process. The review will take place in Beijing and Lanzhou.
- Clark will prepare a draft letter analyzing the WDC responses to the Panel's inquiries for data and asking the centers that have not responded to do so within 1 month.
  - Linn and Clark will find a Chinese scientist in the U.S. or Europe to act as an assistant secretary for the review.
  - Linn will ask Ruttenberg to provide the name(s) of Chinese administrator who established the WDCs originally.

## 2007 Anniversary Activities

The Panel decided that 50th anniversary celebrations of the WDC system will be held in 2007. Options include (1) giving a few talks in IUGG symposia (planning will begin in the fall), (2) scheduling a social event at IUGG, (3) organizing a data symposium in coordination with other data sessions at IUGG, and (4) holding an All WDC conference. The social event was discarded because of cost. The Panel favored hosting an All WDC conference to discuss technical issues and IUGG talks to discuss user issues. A 2-day All WDC conference preceding or following the IUGG symposium would reduce travel costs, but this would make a long meeting even longer, there is no way to arrange letters of invitation, and there would be no local support or financial contribution. The Panel decided to hold the All WDC conference at a different time and place from IUGG. The Chinese offered to host several years ago and they are very good with logistical support, but other centers may also wish to host.

The Panel decided to submit at least 1 data-oriented paper to IGY sessions at IUGG. An All WDC conference will be held at a WDC in 2007. The conference will be about 3 days long, including a Panel meeting.

- Tyupkin and Clark will coordinate participation of WDC presentations in as many IGY association symposia as appropriate.
- Linn will ask for WDC volunteers to host the All WDC conference.

## Panel Incorporation

The WDC Panel was required by ICSU to incorporate and it did so in Boulder, Colorado (USA), because that is where the treasurer is located. The Panel remains an ICSU body and is responsible for providing activity reports upon request and audited financial reports annually.

- Chinnery will send the audited Panel accounts to the ICSU secretariat directly, with a copy to Linn.

## Constitution

The Panel approved a number of changes to the constitution. For example, the provisions were divided into constitution and bylaws. The constitution was changed to remove ICSU as the funding body. The bylaws were changed to eliminate the elected WDC director position, which has not always proved effective. However, WDC directors can serve on the Panel.

- Clark will ask the webmaster to remove the old constitution from the WDC website.
- Linn will send the revised constitution to ICSU by April, along with the proposed new Panel membership.

## Panel Membership

The terms of membership are 4 years, extendable once. Most members have been on the Panel for at least 8 years. Replacing the entire Panel at once would be counterproductive, given the

important activities the Panel has initiated. The Panel decided to replace members on the following schedule:

- 2005: Gvishiani, Linn, and Tyupkin
- 2006: Webster (acting as past president in 2005) and Araki
- 2007: Sobolev
- 2008: Sun Honglie
- 2009: Chinnery

Tyupkin will become director of the Russia WDC coordination office, effective immediately. Linn will continue to serve as director of the USA WDC coordination office.

Efforts will be made to diversify the Panel membership by having more women, young scientists, and scientists in other disciplines, and by improving the geographic distribution (e.g., Europe, developing countries). Sources of names include the ICSU global change programs and IPY.

Smith wondered why ICSU bodies are not represented on the Panel. Linn said that in the past representatives assigned by their organizations did not always participate because they were not interested and/or there were no travel funds. So the constitution was changed in 1996 to allow ICSU bodies to name corresponding members. Only a few groups (CODATA, FAGS, ICSI, URSI, and IUGG) have chosen to have corresponding members to the Panel.

Smith said that the PAA on data and information did not make recommendations about the WDC Panel, but that Panel membership has to change to carry out the PAA recommendations and to bring it in line with other ICSU committees. The new Panel chair will probably also be invited onto the proposed ICSU ad hoc committee to develop a framework for data and information. Therefore, the new Panel membership, particularly the nominations for chair, should be sent to ICSU by the end of May.

The Panel membership will consist of the chair, chair-elect (1-year term before the chair rotates off), vice-chair, secretary, treasurer, 5 representatives of regions, and 3 at-large members. Assistant secretaries will be appointed by the chair as needed on an hoc basis. Assistant secretaries and directors of coordination offices will be invited to participate in Panel meetings as needed.

- Linn (chair), Gvishiani, and Smith will seek candidates for chair-elect, vice-chair, secretary, European Union representative, and 3 at-large members, including 1 WDC director.
- Linn will ask ICSU bodies to nominate corresponding members to the Panel.

## **Finances**

All income for the Panel comes from a 15% charge for managing National Geophysical Data Center (USA) projects. Income from this source varies from year to year and is likely to be less than \$10,000 in 2006. ICSU is no longer a source of operational funds. The cost of running the secretariat is obtained by Linn separately. The total amount available for Panel activities is \$78,000 minus about \$15,000 for this meeting.

The Panel needs operational funds for its activities. A membership fee of \$1,000 per WDC was approved at the 2001 Panel meeting. Subsequent informal discussions suggest that data centers can afford this fee, although a fee may give countries an incentive to consolidate WDCs. Moreover, there may be logistical difficulties in getting money to the Panel. For example, Sobolev said it is illegal for the National Geophysical Committee to pay a WDC fee—Russia has to find a different source of funds. Huang said the China WDC might have a problem paying the fee in U.S. dollars—an official agreement with the government will have to be negotiated.

The Panel considered charging less to associate centers, but this would provide a disincentive for them to advance to full WDC status. The Panel also considered charging less to smaller/poorer data centers, but decided to have a single fixed fee, while not completely excluding some flexibility in applying this charge on a case-by-case basis.

- Chinnery will draft a letter to the WDC directors explaining the need for operational funds and asking for \$1,000 funding per center for 2006. The letter should mention the PAA on data and information and recertification.

### **Partner Centers**

Partner centers have been established in Kenya and Argentina to provide data to developing countries and a means for local scientists to contribute data to global networks. The centers are not necessarily expected to grow into full WDCs, although some may. However, efforts to expand the program have stalled. ICSU gave last year's proposal to add 5 partner centers an "A" rating, but didn't have sufficient funds to fund it. Linn is trying to obtain U.S. Academy of Science funding to hold a meeting of U.S. and European WDC directors and potential South and Central American partners. Other potential partners might be found in Azerbaijan, Chile, and southeast Asia.

Regional centers might also be created, possibly at full WDC status.

- The Panel decided to invite partners or potential partners to the next All WDC conference.
- Grobe will approach his Chilean contact about establishing a virtual WDC for El Niño.
  - Diepenbroek will talk with his contacts at JAMSTEC (Japan) about creating a WDC for IODE.
  - Linn will continue to pursue funding for a meeting with potential Latin America partners.

### **ICSU Proposal**

The amount of money available through ICSU's grant program has greatly declined (to \$350,000) since the U.S. rejoined UNESCO (State Department funding is now sent there instead of ICSU). Smith thought the partner site proposal could be resubmitted with a smaller budget and greater emphasis on capacity building and the PAA on data and information. A regional spin might make the proposal attractive to UNESCO. Consider partnering with a union and a national member in the relevant country.

- Linn and Webster will revise last year's proposal and look for relevant partners, if it can be done quickly.

### **Other Issues**

The WDC MARE is working to assign digital object identifiers (DOI) to datasets. Unlike the system being used commercially, this is a new, nonprofit system. Perhaps the Panel should encourage this more broadly.

Sobolev asked how the WDC system should interact with GRID and other advanced technologies. Webster said that technical issues like state-of-the-art or standards are normally discussed at All WDC conferences. The Panel doesn't have the technical competence to impose them from the top. The Panel tried establishing a technical working group, but there is no money for meetings and work doesn't progress by email.

The Panel will consider drafting a quarterly newsletters to inform WDCs what is happening in the system, posting a "what's new?" column on the website, and/or producing a tri-fold brochure with history, accomplishments, and future plans to help publicize the WDC System.